Server Directed Collective I/O for Arrays in Large Scale Applications

Prof. Marianne Winslett

Subramaniam Mahesh

Department of Computer Science

University of Illinois

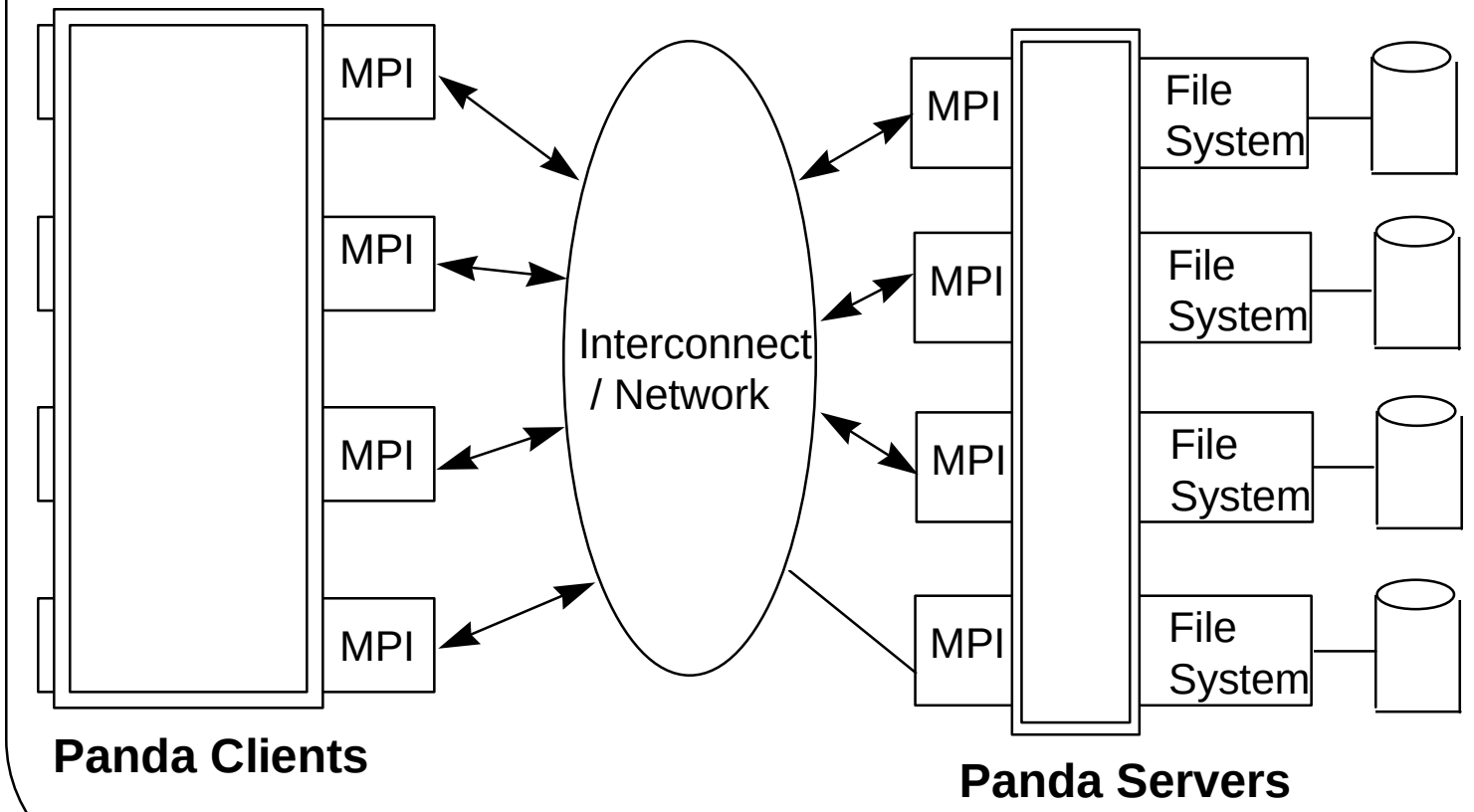
Outline

- **Goals**
- **Server Directed I/O architecture**
- **Performance results**
- **Future Work**
- **Conclusions**

Goals

- **Simple high level interfaces**
  - **Easy to use**
  - **Application portability**
  - **Flexible, efficient underlying implementation**
- **High Performance**
- **Target architectures**
  - **Parallel platforms**
  - **Sequential platforms**

System architecture



Motivation

- **Traditional file systems**
  - **random disk seeks**
  - **prefetching and buffering errors**
- **Server directed I/O**
  - **Servers direct the I/O requests**
  - **sequential reads and writes whenever possible**
  - **minimize random disk seeks**
  - **minimize buffering and prefetching errors**

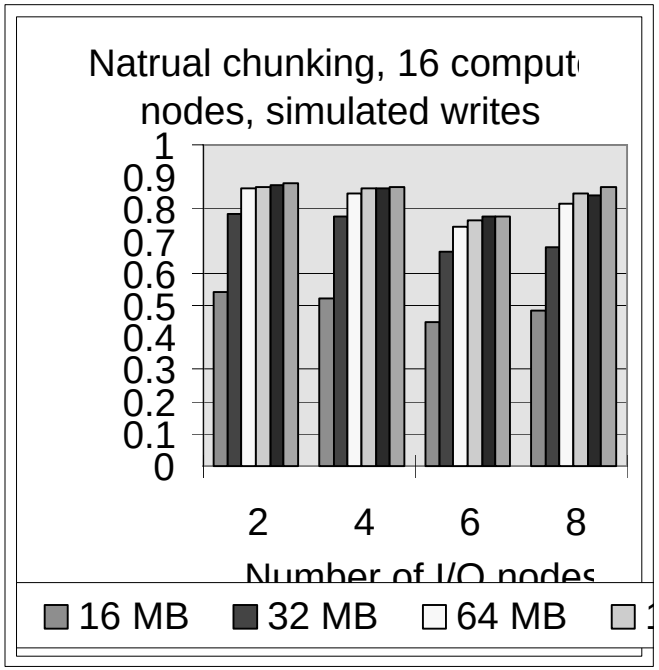
Simple interface

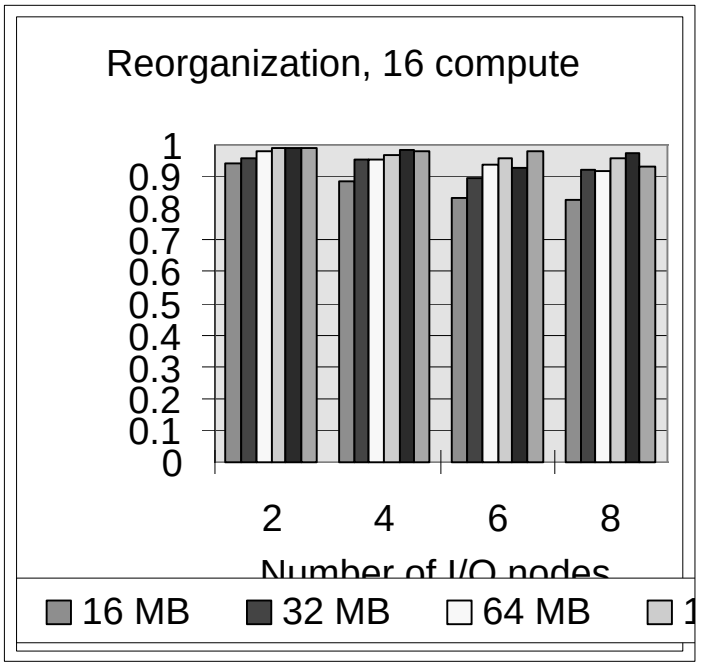
```
Array *density = new Array(“density”, array_rank,
                           array_size, int_size_in_bytes,
                           mem, mem_dis, disk, disk_dist);
Array *temperature = .....
ArrayGRoup *sim = new ArrayGroup(“Sim”, “sim.schema”);
sim->include(temperature);
sim->include(density);
for (int i=0; i<100; i++) {
   compute_next_timestep();
   sim->timestep();
 }
```

Performance measurements

- **Platforms tested**
  - **IBM SP2 (NAS Ames, Cornell Theory Center)**
  - **HP Cluster (NOW)**
- **Parameters**
  - **File system utilization**
  - **Interconnection network utilization**
  - **Scalability**

Performance results





Performance results

- **High performance for large arrays**
- **H3espresso (Black hole application)**
- **Potential for improvement**
  - **finely distributed small arrays**
  - **slower network**

Future Work

- **Performance modeling**
- **Networks of workstations**
- **Shared and Part-time I/O nodes**
- **Application studies**
- **Out-of-core applications**

Conclusions

- **Simple high level interfaces**
  - **easy to use**
  - **flexible and efficient underlying implementation**
  - **application portability**
- **Server Directed I/O**
  - **high performance**